Journal of Multivariate Analysis 100 (2009) 2214–2223

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Regression analysis of multivariate recurrent event data with time-varying covariate effects

Liuquan Sun [a], Liang Zhu [b,*], Jianguo Sun [c]

[a] Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences Beijing 100080, China
[b] Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, 38105, USA
[c] Department of Statistics, University of Missouri-Columbia, Columbia, MO, 65203, USA

ARTICLE INFO

Article history:
Received 1 April 2008
Available online 15 August 2009

AMS 2000 subject classifications:
Primary 62N02
Secondary 62G20

Keywords:
Event history study
Marginal models
Recurrent event data
Time-varying coefficients

ABSTRACT

Recurrent event data occur in many fields and many approaches have been proposed for their analyses (Andersen et al. (1993) [1]; Cook and Lawless (2007) [3]). However, most of the available methods allow only time-independent covariate effects, and sometimes this may not be true. In this paper, we consider regression analysis of multivariate recurrent event data in which some covariate effects may be time-dependent. For the problem, we employ the marginal modeling approach and, especially, estimating equation-based inference procedures are developed. Both asymptotic and finite-sample properties of the proposed estimates are established and an illustrative example is provided.

1. Introduction

Recurrent event data usually refer to data in which the event of interest can occur repeatedly and they arise in many fields such as longitudinal studies, reliability experiments and sociological studies (Andersen et al. [1]; Cook and Lawless [2,3]; Cook et al. [4]). Examples include repeated transient ischemic attacks in cerebrovascular disease study, recurrent pulmonary exacerbations in cystic fibrosis trials, and late infections in bone marrow transplantation. Other examples of recurrent events that often occur in practice include hospitalizations and tumor metastases. Many authors have discussed the analysis of recurrent event data, and they include Andersen and Gill [5], Lawless and Nadeau [6], Lin et al. [7], Lin et al. [8] and Prentice et al. [9].

Sometimes there may exist several related recurrent events, and one is interested in studying them together. For example, transient ischemic attacks may be classified according to location in cardiovascular trials, pulmonary exacerbations may be differentiated by severity, and infections in bone marrow transplantation may be subtyped as bacterial, fungal and viral origins. For the analysis of these multivariate recurrent event data, several inference procedures have also been developed in the literature (Abu-Libdeh et al. [10]; Cai and Schaubel [11]; Clegg et al. [12]; Schaubel and Cai [13]; Spiekerman and Lin [14]). In particular, Cook and Lawless [3] provided a comprehensive review of the existing literature about recurrent events.

For all the methods described above, one limitation is that they were developed only for the situation where covariate effects are constant or time-independent. In practice, however, this may not be true. A common example is that treatment effects may take some time to be effective and then gradually disappear after some time. In other words, the effects vary with

* Corresponding author.
E-mail addresses: slq@amt.ac.cn (L. Sun), liang.zhu@stjude.org (L. Zhu), SunJ@missouri.edu (J. Sun).

doi:10.1016/j.jmva.2009.08.001

time or are time-dependent (Cai and Sun [15]; Lin and Ying [16]; Martinussen and Scheike [17]; Martinussen et al. [18]). In this paper, we consider situations where covariate effects may be time-dependent and develop some semiparametric procedures for regression analysis of multivariate recurrent event data. In addition to allow time-dependent covariate effects, the proposed methods provide a nice graphical summary of time dynamics of covariate effects which cannot be obtained by using the existing methods.

The remainder of this paper is organized as follows. We will begin in Section 2 by introducing some notation and describing the models. An estimation procedure is then presented for both finite- and infinite-dimensional parameters. Section 3 discusses the determination of the proposed estimates, and an iterative algorithm is presented. In Section 4, the asymptotic properties of the proposed estimates are established and some robust estimates of the standard errors of the proposed estimates are given. Note that, in practice, sometimes it may be difficult to know which covariates have time-dependent or time-independent effects. To address this, two test procedures are developed in Section 5 for testing the time-dependence of regression coefficients based on the asymptotic behavior of the cumulative regression functions. Results from an extensive simulation study conducted for evaluating the proposed estimates are reported in Section 6, and Section 7 illustrates the methodology by using a set of bivariate recurrent event data arising from a study of patients with end-stage renal disease. Section 8 contains some concluding remarks.

2. Notation and models

Consider a recurrent event study that involves n independent subjects who experience K different types of recurrent event. For each pair (i, k), let $N^*_{ik}(t)$ denote the number of events of type k that have occurred up to time t for subject i and C_{ik} the censoring or stopping time on the type k event for subject i. In other words, $N^*_{ik}(t)$ is observed only up to time C_{ik}, $i = 1, \ldots, n$, $k = 1, \ldots, K$. If $K = 1$, we then have univariate recurrent event data. Suppose that, for subject i, there exist two types of covariate denoted by $X_{ik}(\cdot)$ and $Z_{ik}(\cdot)$, which are p- and q-dimensional vectors, respectively. Here $X_{ik}(\cdot)$ represents covariates whose effects on the occurrence of the recurrent events of interest may vary with time and $Z_{ik}(\cdot)$ denotes covariates that have fixed or time-independent effects. In the following, we assume that C_{ik} is independent of $N^*_{ik}(\cdot)$ given $X_{ik}(\cdot)$ and $Z_{ik}(\cdot)$. Define $N_{ik}(t) = N^*_{ik}(t \wedge C_{ik})$ and $Y_{ik}(t) = I(C_{ik} \geq t)$, where $a \wedge b = \min(a, b)$ and $I(\cdot)$ denotes the indicator function. Then the observable data consist of $\{N_{ik}(\cdot), Y_{ik}(\cdot), X_{ik}(\cdot), Z_{ik}(\cdot); i = 1, \ldots, n, k = 1, \ldots, K\}$.

To characterize the covariate effects on $N^*_{ik}(t)$, for the kth type events, we assume

$$E\{dN^*_{ik}(t)|X_{ik}(t), Z_{ik}(t)\} = \exp\{\beta_0(t)^T X_{ik}(t) + \gamma_0^T Z_{ik}(t)\}\, d\mu_{0k}(t), \quad k = 1, \ldots, K. \tag{1}$$

Here $dN^*_{ik}(t)$ denotes the increment $N^*_{ik}\{(t+dt)-\} - N^*_{ik}(t-)$ of $N^*_{ik}(t)$ over the small interval $[t, t+dt)$, $\beta_0(t)$ is an unknown p-dimensional vector of time-varying regression coefficients, γ_0 is a q-dimensional vector of unknown regression coefficients, $\mu_{0k}(t)$ is an unspecified baseline mean function, and $d\mu_{0k}(t)$ is the differential of $\mu_{0k}(t)$. Note that here we only specify the marginal behavior of the counting processes N^*_{ik} and leave their relationship arbitrary. Similar models have been discussed by, for example, Cai and Schaubel [11] and Lin et al. [7] when there exist only time-independent regression coefficients. Also note that, without loss of generality, we assume that the regression coefficients are the same for different types of event. If they are type-specific, the methodology developed below still applies by defining a large and new vector of covariates. For example, assume that $\beta_{k0}(t)$ and γ_{k0} differ for different k. In this case, define $\beta_0(t) = (\beta_{10}(t)^T, \ldots, \beta_{K0}(t)^T)^T$, $\gamma_0 = (\gamma_{10}, \ldots, \gamma_{K0})^T$, $X^*_{ik}(t) = (0, \ldots, X_{ik}(t)^T, \ldots, 0)^T$, and $Z^*_{ik}(t) = (0, \ldots, Z_{ik}(t)^T, \ldots, 0)^T$. Then $\beta_{k0}(t)^T X_{ik}(t) + \gamma_{k0}^T Z_{ik}(t)$ can be rewritten as $\beta_0(t)^T X^*_{ik}(t) + \gamma_0^T Z^*_{ik}(t)$ in model (1).

For inference about model (1), in the following, we will focus on estimation of the cumulative regression functions $B_0(t) = \int_0^t \beta_0(s)\, ds$ and γ_0. This is because $B_0(t)$ can be estimated at $n^{1/2}$-rate and further lead to a uniform asymptotic description of the estimators which is needed when one wishes to examine hypotheses about $\beta_0(\cdot)$ (Scheike and Martinussen [19]). Also, the available information about $\beta_0(t)$ at any particular time point is in general too limited to yield a consistent estimate directly. If $\beta_0(t)$ is of interest, one can develop some smooth estimates by using an estimate of $B_0(t)$ such as kernel estimates.

To present the estimation procedure, let $N_k(t) = (N_{1k}(t), \ldots, N_{nk}(t))^T$, $N\cdot_k(t) = \frac{1}{n}\sum_{i=1}^n N_{ik}(t)$, $Y_k(t) = (Y_{1k}(t), \ldots, Y_{nk}(t))^T$, $X_k(t) = (X_{1k}(t), \ldots, X_{nk}(t))^T$, $Z_k(t) = (Z_{1k}(t), \ldots, Z_{nk}(t))^T$, for $k = 1, \ldots, K$. Define

$$\phi_{ik}(t; \beta, \gamma) = Y_{ik}(t) \exp\{\beta(t)^T X_{ik}(t) + \gamma^T Z_{ik}(t)\},$$
$$\phi_k(t; \beta, \gamma) = (\phi_{1k}(t; \beta, \gamma), \ldots, \phi_{nk}(t; \beta, \gamma))^T,$$
$$\Phi_k(t; \beta, \gamma) = \mathrm{diag}\{\phi_{1k}(t; \beta, \gamma), \ldots, \phi_{nk}(t; \beta, \gamma)\},$$
$$S_k^{(0)}(t; \beta, \gamma) = n^{-1} \sum_{i=1}^n \phi_{ik}(t; \beta, \gamma),$$
$$S_k^{(x)}(t; \beta, \gamma) = n^{-1} \sum_{i=1}^n \phi_{ik}(t; \beta, \gamma) X_{ik}(t),$$
$$S_k^{(z)}(t; \beta, \gamma) = n^{-1} \sum_{i=1}^n \phi_{ik}(t; \beta, \gamma) Z_{ik}(t),$$

$$E_k^{(x)}(t;\beta,\gamma)=S_k^{(x)}(t;\beta,\gamma)/S_k^{(0)}(t;\beta,\gamma),$$

and $E_k^{(z)}(t;\beta,\gamma)=S_k^{(z)}(t;\beta,\gamma)/S_k^{(0)}(t;\beta,\gamma)$. Also let $\bar{X}_k(t;\beta,\gamma)$ be the $n\times p$ matrix with rows $E_k^{(x)}(t;\beta,\gamma)^T$, $\bar{Z}_k(t;\beta,\gamma)$ the $n\times q$ matrix with rows $E_k^{(z)}(t;\beta,\gamma)^T$, and

$$M_{ik}(t)=N_{ik}(t)-\int_0^t Y_{ik}(s)\exp\{\beta_0(s)^T X_{ik}(s)+\gamma_0^T Z_{ik}(s)\}\mathrm{d}\mu_{0k}(s),$$

for $k=1,\ldots,K$ and $i=1,\ldots,n$. It can be easily seen that under model (1), the $M_{ik}(t)$'s are zero-mean stochastic processes. It follows that, for given $\beta(t)$ and γ, a natural estimate of $\mu_{0k}(t)$ is given by the solution to

$$\sum_{i=1}^n\Big[\mathrm{d}N_{ik}(t)-Y_{ik}(t)\exp\{\beta(t)^T X_{ik}(t)+\gamma^T Z_{ik}(t)\}\mathrm{d}\mu_{0k}(t)\Big]=0,\quad 0\le t\le\tau,\ k=1,\ldots,K,$$

where τ is a prespecified constant such that $P(C_{ik}\ge\tau)>0$. Let $\hat{\mu}_{0k}(t;\beta,\gamma)$ denote the estimate defined above. Then we have

$$\hat{\mu}_{0k}(t;\beta,\gamma)=\int_0^t\frac{\mathrm{d}N_{\cdot k}(u)}{S_k^{(0)}(u;\beta,\gamma)}. \tag{2}$$

For estimation of $\beta_0(t)$ and γ_0, following the idea behind the generalized estimating equation approach (Liang and Zeger [20]) and (1) and (2), we propose the following estimating equations:

$$\sum_{k=1}^K X_k(t)^T\left[\mathrm{d}N_k(t)-\frac{\phi_k(t;\beta,\gamma)\mathrm{d}N_{\cdot k}(t)}{S_k^{(0)}(t;\beta,\gamma)}\right]=0,\quad 0\le t\le\tau,$$

and

$$\sum_{k=1}^K\int_0^\tau Z_k(t)^T\left[\mathrm{d}N_k(t)-\frac{\phi_k(t;\beta,\gamma)\mathrm{d}N_{\cdot k}(t)}{S_k^{(0)}(t;\beta,\gamma)}\right]=0.$$

It is apparent that solving the equations given above is not straightforward. For this, following Scheike and Martinussen [19], we propose to use the iterative algorithm given in the next section.

3. An iterative algorithm

To describe the algorithm, let $\hat{B}^{(l)}(t)$ and $\hat{\gamma}^{(l)}$ denote the estimates of $B_0(t)$ and γ_0 obtained in the lth iteration and $\hat{\beta}^{(l)}(t)$ a smooth estimate of $\beta_0(t)$ based on $\hat{B}^{(l)}(t)$, which will be discussed later. Note that the application of the Taylor expansion to $\phi_k(t;\beta,\gamma)S_k^{(0)}(t;\beta,\gamma)^{-1}$ around $\hat{\beta}^{(l)}(t)$ and $\hat{\gamma}^{(l)}$ gives the updating equations:

$$\begin{aligned}\sum_{k=1}^K E_{xx}^{(k)(l)}(t)\left\{\hat{\beta}^{(l+1)}(t)-\hat{\beta}^{(l)}(t)\right\}\frac{\mathrm{d}N_{\cdot k}(t)}{S_k^{(0)(l)}(t)}&=\frac{1}{n}\sum_{k=1}^K\{X_k(t)-\bar{X}_k^{(l)}(t)\}^T\\ &\times\left[\mathrm{d}N_k(t)-\Phi_k^{(l)}(t)\{Z(t)-\bar{Z}^{(l)}(t)\}(\hat{\gamma}^{(l+1)}-\hat{\gamma}^{(l)})\frac{\mathrm{d}N_{\cdot k}(t)}{S_k^{(0)(l)}(t)}\right],\end{aligned} \tag{3}$$

and

$$\begin{aligned}&\frac{1}{n}\sum_{k=1}^K\int_0^\tau\{Z_k(t)-\bar{Z}_k^{(l)}(t)\}^T\mathrm{d}N_k(t)-\sum_{k=1}^K\int_0^\tau E_{zx}^{(k)(l)}(t)\{\hat{\beta}^{(l+1)}(t)-\hat{\beta}^{(l)}(t)\}\frac{\mathrm{d}N_{\cdot k}(t)}{S_k^{(0)(l)}(t)}\\ &=\sum_{k=1}^K\int_0^\tau E_{zz}^{(k)(l)}(t)\left(\hat{\gamma}^{(l+1)}-\hat{\gamma}^{(l)}\right)\frac{\mathrm{d}N_{\cdot k}(t)}{S_k^{(0)(l)}(t)}.\end{aligned} \tag{4}$$

In these expressions, for $k=1,\ldots,K$,

$$\begin{aligned}E_{xx}^{(k)}(t)&=n^{-1}\{X_k(t)-\bar{X}_k(t;\beta,\gamma)\}^T\Phi_k(t;\beta,\gamma)\{X_k(t)-\bar{X}_k(t;\beta,\gamma)\},\\ E_{zx}^{(k)}(t)&=n^{-1}\{Z_k(t)-\bar{Z}_k(t;\beta,\gamma)\}^T\Phi_k(t;\beta,\gamma)\{X_k(t)-\bar{X}_k(t;\beta,\gamma)\},\\ E_{zz}^{(k)}(t)&=n^{-1}\{Z_k(t)-\bar{Z}_k(t;\beta,\gamma)\}^T\Phi_k(t;\beta,\gamma)\{Z_k(t)-\bar{Z}_k(t;\beta,\gamma)\},\end{aligned}$$

and $S_k^{(0)(l)}(t)$, $\Phi_k^{(l)}(t)$, $\bar{X}_k^{(l)}(t)$, $\bar{Z}_k^{(l)}(t)$, $E_{xx}^{(k)(l)}(t)$, $E_{zx}^{(k)(l)}(t)$ and $E_{zz}^{(k)(l)}(t)$ denote the values of these quantities at $\beta(t)=\hat{\beta}^{(l)}(t)$ and $\gamma=\hat{\gamma}^{(l)}$.

Let $\hat{\lambda}_{0k}^{(l)}(t)$ denote a smooth estimate of $\lambda_{0k}(t) = \mathrm{d}\mu_{0k}(t)/\mathrm{d}t$ based on $\hat{\mu}_{0k}(t; \hat{\beta}^{(l)}, \hat{\gamma}^{(l)})$ (see below). Define

$$A_z^{(l)}(t) = \sum_{k=1}^{K} E_{zx}^{(k)(l)}(t)\hat{\lambda}_{0k}^{(l)}(t), \qquad A_x^{(l)}(t) = \sum_{k=1}^{K} E_{xx}^{(k)(l)}(t)\hat{\lambda}_{0k}^{(l)}(t),$$

and

$$D^{(l)}(\tau) = \sum_{k=1}^{K}\int_0^{\tau}\left[E_{zz}^{(k)(l)}(t) - A_z^{(l)}(t)A_x^{(l)}(t)^{-1}E_{zx}^{(k)(l)}(t)^T\right]\frac{\mathrm{d}N\cdot_k(t)}{S_k^{(0)(l)}(t)}.$$

Then by replacing $\mathrm{d}N\cdot_k(t)/S_k^{(0)(l)}(t)$ with $\hat{\lambda}_{0k}^{(l)}(t)\mathrm{d}t$, one can derive from (3) and (4) the updated estimate of γ as

$$\hat{\gamma}^{(l+1)} = H_\gamma(\hat{\gamma}^{(l)}) = \hat{\gamma}^{(l)} + \frac{1}{n}D^{(l)}(\tau)^{-1} \times \sum_{k=1}^{K}\int_0^{\tau}\left[\{Z_k(t) - \bar{Z}_k^{(l)}(t)\}^T - A_z^{(l)}(t)A_x^{(l)}(t)^{-1}\{X_k(t) - \bar{X}_k^{(l)}(t)\}^T\right]\mathrm{d}N_k(t). \tag{5}$$

Given $\hat{\gamma}^{(l+1)}$, we have

$$\hat{B}^{(l+1)}(t) = H_b(\hat{B}^{(l)})(t) = \int_0^t \hat{\beta}^{(l)}(u)\mathrm{d}u + n^{-1}\sum_{k=1}^{K}\int_0^t A_x^{(l)}(u)^{-1}\{X_k(u) - \bar{X}_k^{(l)}(u)\}^T \times\left[\mathrm{d}N_k(u) - \Phi_k^{(l)}(u)\{Z_k(u) - \bar{Z}_k^{(l)}(u)\}(\hat{\gamma}^{(l+1)} - \hat{\gamma}^{(l)})\frac{\mathrm{d}N\cdot_k(u)}{S_k^{(0)(l)}(u)}\right]. \tag{6}$$

In the algorithm described above, we need smooth estimates of $\lambda_{0k}(t)$ and $\beta_0(t)$ and many types of such estimates can be used. In the following, we will use the kernel estimates given by

$$\hat{\lambda}_{0k}(t) = \int h^{-1}G\left(\frac{u-t}{h}\right)\mathrm{d}\hat{\mu}_{0k}(u), \quad k = 1, \ldots, K \tag{7a}$$

and

$$\hat{\beta}(t) = \int h^{-1}G\left(\frac{u-t}{h}\right)\mathrm{d}\hat{B}(u), \tag{7b}$$

where G is a symmetric kernel function with a compact support and h denotes the bandwidth. Among others, Scheike and Martinussen [19] used these estimates.

Given the description above, the iterative algorithm can be summarized as follows.

Step 0. Choose initial estimates $\hat{\beta}^{(0)}(t)$ and $\hat{\gamma}^{(0)}$.

Step 1. At the lth iteration, obtain $\hat{\mu}_{0k}^{(l)}$ by (2) and then the kernel estimate $\hat{\lambda}_{0k}^{(l)}$ by (7a), for $k = 1, \ldots, K$.

Step 2. Calculate the updated estimates $\hat{\gamma}^{(l+1)}$ and $\hat{B}^{(l+1)}$ using Eqs. (5) and (6), respectively, and then the kernel estimate $\hat{\beta}^{(l+1)}$ by (7b).

Step 3. Return to step 1 for the $(l+1)$th iteration until convergence.

Note that many choices can be used for the initial estimates of $\hat{\beta}^{(0)}(t)$ and $\hat{\gamma}^{(0)}$. One simple choice is to assume that $\beta_0(t)$ is time-independent and then to use the estimates given in Lin et al. [7]. For the convergence, also several criteria can be applied, and in the numerical studies below, we used the absolute differences between the iterative estimates of the parameters. The algorithm converges most times in general, but non-convergence could occur occasionally depending on the set-ups. In general, the speed is around 98 s for one run with $n = 200$ in Matlab.

4. Asymptotic properties

In this section, we will establish the asymptotic properties of the estimates defined in the previous sections. Let $\hat{B}(t)$ and $\hat{\gamma}$ denote the estimates of $B_0(t)$ and γ_0 and $\hat{\beta}(t)$, $\hat{\mu}_{0k}(t)$ and $\hat{\lambda}_{0k}(t)$ the resulting estimates of $\beta_0(t)$, $\mu_{0k}(t)$ and $\lambda_{0k}(t)$, respectively. Also, let $\hat{A}_z(t)$, $\hat{A}_x(t)$, $\hat{D}(\tau)$, $\hat{M}_{ik}(t)$ and $\hat{S}^{(0)}(t)$ denote the quantities defined in the previous sections with all unknown parameters replaced by their estimates. In the following, we first establish the existence of $\hat{B}(t)$ and $\hat{\gamma}$ and their asymptotic properties and then describe the asymptotic behavior of $\hat{\mu}_{0k}(t)$. The proofs of the results will be deferred to the supplemental technical report [21].

Theorem 1. *Assume that the regularity conditions* (C1)–(C5) *stated in* Appendix A *hold; then we have*

(i) *with probability tending to one,* (5) *and* (6) *have unique solutions* $H_\gamma(\hat\gamma) = \hat\gamma$ *and* $H_b(\hat B) = \hat B$ *such that* $\|\hat\gamma - \gamma_0\| = O_p(n^{-1/2})$ *and* $\sup_{0\le t\le\tau} \|\hat B(t) - B_0(t)\| = O_p(n^{-1/2})$, *where* $\|v\| = (v^T v)^{1/2}$ *for a vector* v;

(ii) $n^{1/2}\{\hat\gamma - \gamma_0\}$ *is asymptotically normally distributed with zero mean and covariance matrix that can be consistently estimated by* $\hat\Sigma = n^{-1}\sum_{i=1}^n \hat\xi_i(\tau)\hat\xi_i(\tau)^T$, *where*

$$\hat\xi_i(\tau) = \hat D(\tau)^{-1}\sum_{k=1}^K \int_0^\tau \Big[\{Z_{ik}(t) - \hat E_k^{(z)}(t)\} - \hat A_z(t)\hat A_x(t)^{-1}\{X_{ik}(t) - \hat E_k^{(x)}(t)\}\Big]\mathrm{d}\hat M_{ik}(t); \tag{8}$$

(iii) $n^{1/2}\{\hat B(t) - B_0(t)\}$ *converges weakly to a zero-mean Gaussian process whose covariance function at* (s,t) *can be consistently estimated by* $\hat\Gamma_b(s,t) = n^{-1}\sum_{i=1}^n \hat\eta_i(s)\hat\eta_i(t)^T$, *where*

$$\hat\eta_i(t) = \sum_{k=1}^K \int_0^t \hat A_x(u)^{-1}\{X_{ik}(u) - \hat E_k^{(x)}(u)\}\mathrm{d}\hat M_{ik}(u) - \int_0^t \hat A_x(u)^{-1}\hat A_z(u)^T \mathrm{d}u\, \hat\xi_i(\tau). \tag{9}$$

For a given kernel function $G(t)$, define

$$\upsilon_0 = \int G(u)^2\mathrm{d}u, \qquad \upsilon_1 = \int u^2 G(u)\mathrm{d}u.$$

Let $\beta_0''(t)$ denote the second derivative of $\beta_0(t)$ and $e_k^{(x)}(t)$ the limit of $E_k^{(x)}(t;\beta_0,\gamma_0)$. Then we have the following results for $\hat\mu_{0k}(t)$.

Theorem 2. *Again assume that the regularity conditions* (C1)–(C5) *stated in* Appendix A *hold; then we have*

(i) $\sup_{0\le t\le\tau} |\hat\mu_{0k}(t) - \mu_{0k}(t)| = O_p(n^{-1/4})$, $k = 1,\ldots,K$;

(ii) $n^{1/2}\{\hat\mu_{0k}(t) - \mu_{0k}(t) + \frac{1}{2}\upsilon_1 h^2 \int_0^t e_k^{(x)}(u)\beta_0''(u)\mathrm{d}\mu_{0k}(u)\}$, *for* $k = 1,\ldots,K$, *converges weakly to a zero-mean Gaussian process whose covariance function at* (s,t) *can be consistently estimated by* $\hat\Gamma_k(s,t) = n^{-1}\sum_{i=1}^n \hat\varphi_{ik}(s)\hat\varphi_{ik}(t)$, *where*

$$\hat\varphi_{ik}(t) = \int_0^t \frac{\mathrm{d}\hat M_{ik}(u)}{\hat S_k^{(0)}(u)} - \int_0^t \hat E_k^{(z)}(u)^T\mathrm{d}\hat\mu_{0k}(u)\,\hat\xi_i(\tau) - \int_0^t \hat E_k^{(x)}(u)^T\hat\lambda_{0k}(u)\mathrm{d}\hat\eta_i(u). \tag{10}$$

Sometimes one may be interested in constructing confidence bands for $B_0(t)$ and $\mu_{0k}(t)$, and one way for this is to use the results given above. However, it may be analytically difficult since the limiting Gaussian processes for $n^{1/2}\{\hat B(t) - B_0(t)\}$ and $n^{1/2}\{\hat\mu_{0k}(t) - \mu_{0k}(t)\}$ $(k = 1,\ldots,K)$ do not have independent increments. To this end, we propose to use the simulation approach described in Lin et al. [7] to approximate these limiting distributions. Specifically, let $\{G_i;\, i = 1,\ldots,n\}$ be a simple random sample of size n from the standard normal distribution and independent of the observed data. Then one can construct the simultaneous confidence bands for $B_0(t)$ and $\mu_{0k}(t)$ by replacing $\hat M_{ik}(t)$ with $G_i\hat M_{ik}(t)$ in (9) and (10) and repeatedly generating normal random samples $\{G_i;\, i = 1,\ldots,n\}$ given the observed data.

5. Test of time-varying covariate effects

As mentioned above, in practice, one may not know which covariates or if indeed any of them have time-dependent or time-varying effects. To address this, one way is to test a hypothesis such as

$$H_0 : \beta_{0j}(t) \quad \text{is a constant} \tag{11}$$

for the jth component of $\beta_0(t)$. In the following, two procedures will be presented for this and we will then discuss testing $B_0(t) = 0$. Of course, one may be interested in testing $\gamma_0 = 0$, and for this, one can simply apply the Wald test.

For given j, let $B_{0j}(t)$ denote the jth element of $B_0(t)$ and define $V(t) = B_{0j}(t) - B_{0j}(\tau)t/\tau$ and $\hat V(t) = \hat B_j(t) - \hat B_j(\tau)t/\tau$. Then under the hypothesis H_0, we have $V(t) \equiv 0$, and this suggests the following two test statistics for (11). One is the Kolmogorov–Smirnov type statistic defined as

$$\mathcal{F}_1 = \sup_{0\le t\le\tau}\left|n^{1/2}\hat V(t)\right|$$

and the other is the Cramér–von Mises type statistic given by

$$\mathcal{F}_2 = \int_0^\tau n\hat V(t)^2\mathrm{d}t.$$

One will reject H_0 if one or both statistics are away from zero.

To apply $\mathcal{F}_1$ and $\mathcal{F}_2$, we need their distributions. For this, note that it follows from Theorem 1(ii) that under H_0 the asymptotic distribution of $n^{1/2}\{\hat{V}(t) - V(t)\}$ is asymptotically equivalent to the zero-mean Gaussian process

$$\hat{W}(t) = n^{-1/2} \sum_{i=1}^{n} \left\{ \hat{\eta}_{ij}(t) - \hat{\eta}_{ij}(\tau)t/\tau \right\},$$

where $\hat{\eta}_{ij}(t)$ is the jth element of $\hat{\eta}_i(t)$ defined in (9). Furthermore, following Lin et al. [7], one can show that the distribution of the process $\hat{W}(t)$ can be approximated by that of the zero-mean Gaussian process

$$\widetilde{W}(t) = n^{-1/2} \sum_{i=1}^{n} \left\{ \hat{\eta}_{ij}(t) - \hat{\eta}_{ij}(\tau)t/\tau \right\} G_i,$$

where $(G_1, \ldots, G_n)$ are independent standard normal variables independent of the observed data. Thus one can approximate the distributions of $\mathcal{F}_1$ and $\mathcal{F}_2$ by those of

$$\tilde{\mathcal{F}}_1 = \sup_{0 \le t \le \tau} \left| \widetilde{W}(t) \right|$$

and

$$\tilde{\mathcal{F}}_2 = \int_0^{\tau} \widetilde{W}(t)^2 \mathrm{d}t,$$

respectively, which can be obtained by a large number of their realizations generated by repeatedly generating the normal random sample $(G_1, \ldots, G_n)$ while fixing the observed data. As a consequence, one can determine the approximate critical values of the tests of the hypothesis H_0 based on the statistics $\mathcal{F}_1$ or $\mathcal{F}_2$.

To test $H_0 : B_{0j}(t) \equiv 0$, similarly, one can consider the statistic

$$\mathcal{F}_3 = \sup_{0 \le t \le \tau} \left| \frac{\hat{B}_j(t)}{\hat{\sigma}_j(t)} \right|$$

and reject the $B_{0j}(t) \equiv 0$ if $\mathcal{F}_3$ is away from zero, where $\hat{\sigma}_j(t)$ is an estimate of the standard error of $\hat{B}_j(t)$. To determine the critical value, again as above, one can approximate the distribution of $\mathcal{F}_3$ using that of

$$\tilde{\mathcal{F}}_3 = \sup_{0 \le t \le \tau} \left| \frac{n^{-1/2} \sum_{i=1}^{n} \hat{\eta}_{ij}(t)\, G_i}{[n^{-1} \sum_{i=1}^{n} \hat{\eta}_{ij}^2(t)]^{1/2}} \right|$$

and by repeatedly generating the normal random sample $(G_1, \ldots, G_n)$.

6. Numerical results

A simulation study was conducted to assess the performance of the estimates of unknown parameters proposed in the previous sections. In the study, we considered the situation where there exist $K = 2$ types of recurrent event of interest and two covariates ($p = q = 1$). One covariate was assumed to have time-varying effect on the occurrence of the events and the effect of the other covariate was supposed to be time-independent. Also, in the study it was assumed that the underlying recurrent event processes $N_{i1}^*(t)$ and $N_{i2}^*(t)$ are mixed Poisson processes with intensity functions

$$\lambda_{i1}(t) = \nu_i \exp(-0.5) \exp[\beta_0(t) X_{i1} + \gamma_0 Z_{i1}]$$

and

$$\lambda_{i2}(t) = \nu_i \exp(-1) \exp[\beta_0(t) X_{i2} + \gamma_0 Z_{i2}],$$

respectively. In these models, the ν_i's are latent variables generated from the gamma distribution with mean one and variance σ^2, and two functions for $\beta_0(t)$ were considered, which are $\beta_1(t) = 0.35 \sin(2t)$ for $0 \le t < \pi/2$ and $\sin(2t)$ for $t > \pi/2$ and $\beta_2(t) = \log(1 + t)$. Furthermore, it was supposed that $X_{i1} = X_{i2}$ with the X_{i1}'s generated from the standard normal distribution or the uniform distribution over $(0, 1)$ and $Z_{i1} = Z_{i2}$ with the Z_{i1}'s following the standard normal distribution. For C_{ik}'s, it was supposed that either $C_{ik} = 3$ with probability p_0 or $C_{ik} = \tau = 5$. For estimation of $\lambda_{0k}(t)$ and $\beta_0(t)$, we used the kernel estimates (7a) and (7b) with the kernel function $G(t) = [\cos(\pi\, t) + 1]/2\, (-1 \le t \le 1)$ and the bandwidth h around 2.5 and 1, respectively.

Table 1 presents the simulation results for the estimation of γ_0 with the true $\gamma_0 = 0.3$, $n = 100$ or 200, $p_0 = 0.25$ or 0.5, and $\sigma^2 = 0$ or 0.25 and based on 1000 replications. For the results with $\beta_0(t) = \beta_1(t)$, both the X_i's and Z_i's were generated from the standard normal distribution, while for the results with $\beta_0(t) = \beta_2(t)$, the X_i's and Z_i's were generated, respectively, from the uniform distribution over $(0, 1)$ and the standard normal distribution. The results include the biases

Table 1
Simulation results for the estimation of γ_0.

$\beta_0(t)$	p_0	σ^2	$n = 100$				$n = 200$			
			Bias	SSD	ESE	CP	Bias	SSD	ESE	CP
$\beta_1(t)$	0.25	0	0.0042	0.0443	0.0436	0.940	−0.0009	0.0315	0.0316	0.946
	0.25	0.25	−0.0055	0.0705	0.0676	0.935	0.0006	0.0494	0.0478	0.951
	0.50	0	0.0000	0.0470	0.0464	0.938	0.0007	0.0330	0.0332	0.944
	0.50	0.25	−0.0055	0.0747	0.0691	0.919	−0.0008	0.0533	0.0499	0.933
$\beta_2(t)$	0.25	0	−0.0027	0.0342	0.0334	0.939	0.0012	0.0243	0.0238	0.951
	0.25	0.25	0.0023	0.0634	0.0610	0.923	0.0007	0.0448	0.0425	0.936
	0.50	0	−0.0018	0.0406	0.0359	0.913	0.0008	0.0249	0.0258	0.956
	0.50	0.25	−0.0004	0.0689	0.0661	0.935	−0.0005	0.0480	0.0461	0.938

Note: SSD represents the sample standard deviation of the estimates, ESE represents the mean of the estimated standard errors, and CP represents the empirical 95% coverage probability.

Table 2
Estimated sizes of the test procedures at the significance level 0.05.

p_0	σ^2	For $\beta_0(t)$ being a constant		For $B_0(t) = 0$
		$\mathcal{F}_1$	$\mathcal{F}_2$	$\mathcal{F}_3$
0.25	0	0.058	0.050	0.048
0.25	0.25	0.052	0.054	0.045
0.50	0	0.054	0.052	0.055
0.50	0.25	0.054	0.056	0.051

Table 3
Estimated powers of the test procedures at the significance level 0.05.

$\beta_0(t)$	p_0	σ^2	$n = 100$			$n = 200$		
			$\beta_0(t)$ is constant		$B_0(t) = 0$	$\beta_0(t)$ is constant		$B_0(t) = 0$
			$\mathcal{F}_1$	$\mathcal{F}_2$	$\mathcal{F}_3$	$\mathcal{F}_1$	$\mathcal{F}_2$	$\mathcal{F}_3$
$\beta_1(t)$	0.25	0	1.000	1.000	0.996	1.000	1.000	1.000
	0.25	0.25	0.999	0.999	0.817	1.000	1.000	0.996
	0.50	0	1.000	1.000	0.985	1.000	1.000	1.000
	0.50	0.25	1.000	1.000	0.896	1.000	1.000	1.000
$\beta_1(t)$	0.25	0	0.921	0.956	1.000	1.000	1.000	1.000
	0.25	0.25	0.935	0.961	0.998	0.998	0.999	1.000
	0.50	0	0.859	0.904	1.000	0.993	0.996	1.000
	0.50	0.25	0.813	0.847	0.988	0.986	0.990	1.000

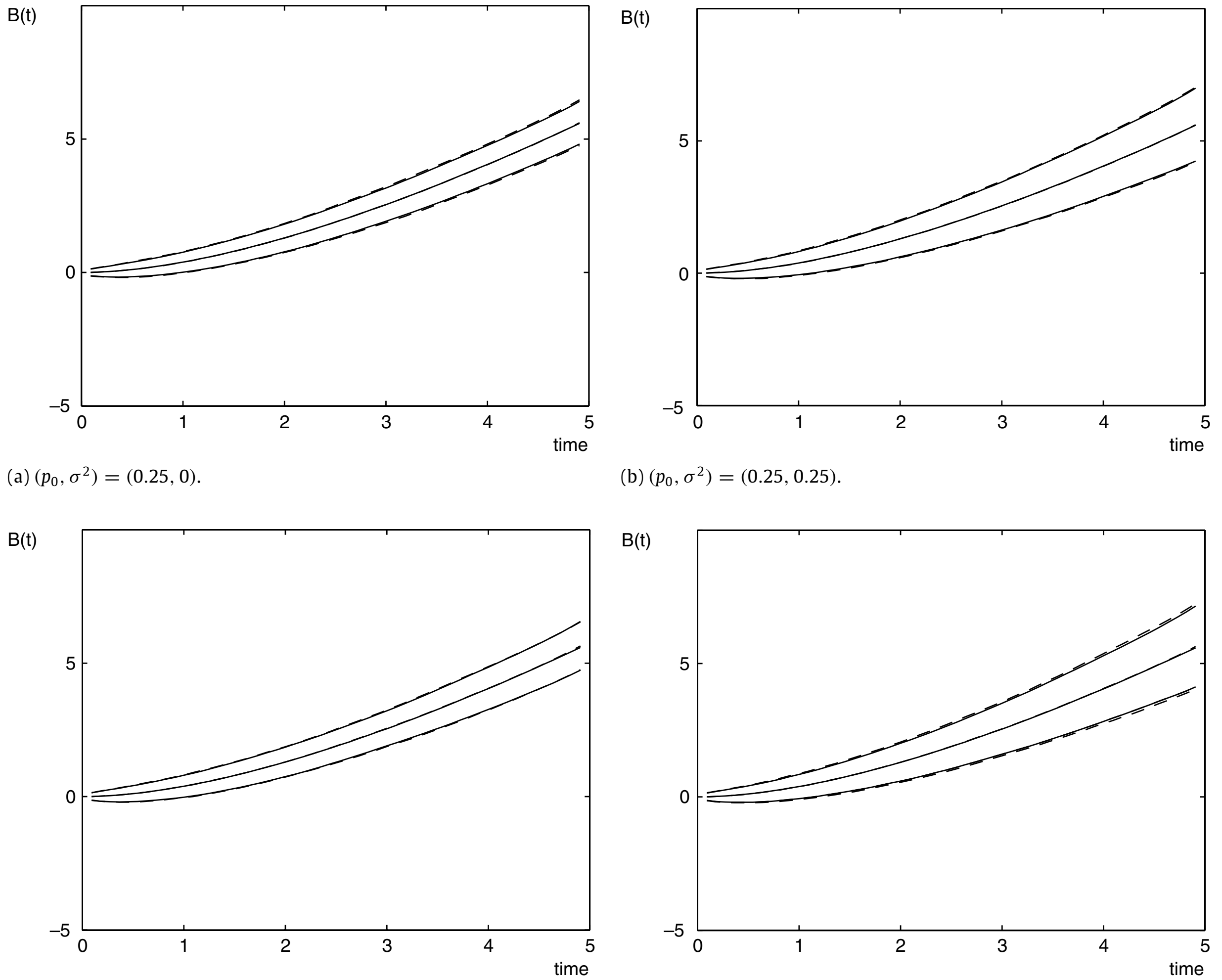
(BIAS) given by the sample means of the point estimates $\hat{\gamma}$ minus the true value, the sample standard deviations of the estimates $\hat{\gamma}$ (SSD), the averages of the estimated standard errors (ESE), and the 95% empirical coverage probabilities (CP). These results suggest that the proposed estimate $\hat{\gamma}$ seems to be unbiased and the given variance estimate also seems to be reasonable, especially for the situations where $n = 200$.

The summary results for estimation of $\beta_0(t)$ are displayed in Fig. 1 for the situations considered in Table 1 in which $n = 200$ and $\beta_0(t) = \beta_2(t)$. The figures (a)–(d) correspond, respectively, to the situations where $(p_0, \sigma^2) = (0.25, 0), (0.25, 0.25), (0.50, 0)$ and $(0.50, 0.25)$. In these figures, we plotted the true $B_0(t)$ and the pointwise 95% confidence bands based on sample standard deviations (solid lines for both) along with the averages of the estimates $\hat{B}(t)$ and the estimated pointwise 95% confidence bands based on the averages of the estimated standard deviations (dashed lines for both). The plots show that both the point and the variance estimates appear to perform well. The plots for the situations where $\beta_0(t) = \beta_1(t)$ are similar and are not presented.

We also investigated the size and power of the three test procedures presented in Section 4 at the significance level of 5%. Table 2 presents the empirical sizes obtained based on 1000 sets of the simulated data for the situations similar to those considered in Table 1 with $n = 200$ and $\beta_0(t) = 0.1$ and 0 for testing $\beta_0(t)$ being a constant and H_0: $B_0(t) = 0$, respectively. The results for the empirical power are given in Table 3 under the same situations as with Table 2. It can be seen from Tables 2 and 3 that all three procedures seem to have correct sizes and to show good power. In particular, the procedure based on $\mathcal{F}_2$ seems to have greater power than that based on $\mathcal{F}_2$, but the difference does not seem to be significant.

7. An application

In this section we apply the methodology proposed in the previous sections to a set of bivariate recurrent event data ($K = 2$) provided by the Canadian Organ Replacement Register of the Canadian Institute for Health Information. The data arose from a study of technical failure (TF) rates among patients who had end-stage renal disease and were receiving



(a) $(p_0, \sigma^2) = (0.25, 0)$.

(b) $(p_0, \sigma^2) = (0.25, 0.25)$.

(c) $(p_0, \sigma^2) = (0.050, 0)$.

(d) $(p_0, \sigma^2) = (0.25, 0.25)$.

Fig. 1. Estimates of the cumulative regression functions with the pointwise 95% confidence bands. Solid lines are true $B(t)$ and sample pointwise confidence bands; dashed lines are estimated $B(t)$ and estimated pointwise confidence bands.

continuous ambulatory peritoneal dialysis (CAPD), a method of dialysis in which a catheter is implemented into the lower abdomen. A TF is considered to occur when the patient is forced to switch from CAPD to another method of dialysis, and there can be several causes for the occurrence of a TF. Also, the occurrences of TF or switches are usually temporary and thus it is reasonable to treat a TF as a recurrent event (Schaubel and Cai [13]). In the following analysis, we will focus on two types of cause for the occurrence of a TF: inadequate dialysis and all others. That is, we have a set of two types of recurrent event data.

The data considered here consist of $n = 5366$ patients who initiated CAPD between 1 January 1990 and 31 December 1998 and were followed up to 31 December 1998. During the study, 1853 TFs were observed in total, and among these, 496 or about 27% were due to inadequate dialysis. The covariates of primary interest are the gender of the patients and the age and diabetes status (yes or no) at the start of CAPD treatment. Knowledge of the effects of these covariates on the TF rates is very useful for nephrologists and health care administrators at the hospitals in which CAPD is practiced.

To analyze the data, we define $N_{i1}^*(t)$ and $N_{i2}^*(t)$ to be the numbers of TFs that had occurred from the start of CAPD treatment up to day t from patient i due to inadequate dialysis and other causes, respectively. Also, we define the covariate corresponding to gender to be 1 for male patients and 0 otherwise and the covariate representing diabetes status to be 1 if the initial status is yes and 0 otherwise. First, we assumed that all three covariates had time-varying effects on the TF rates. The application of the test procedure based on $\mathcal{F}_1$ proposed in the previous sections gave p-values of 0.077, 0.900 and 0.005 for testing the time-dependence of the effects of age, gender and diabetes status, respectively. The test procedure based on $\mathcal{F}_2$ gave similar p-values for gender and diabetes status but a smaller p-value for age. These results and the study of the estimates of the cumulative regression function $B(t)$ suggest that the effects of both age and gender seem to be time-independent. For the results reported above and below, we used the same kernel function as that in Section 5 with the bandwidth equal to 300 and 200 for $\lambda_0(t)$ and $\beta_0(t)$, respectively. We tried several other bandwidths, and the results were similar.

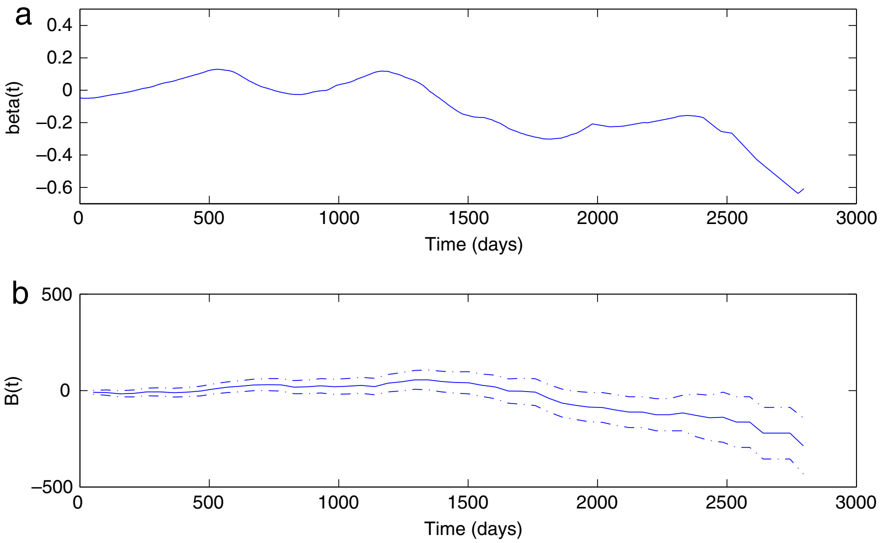


Fig. 2. (a) The estimate of the effect of diabetes status. (b) The estimate of the cumulative effect of diabetes status with its 95% pointwise confidence bands.

Then we assumed that both age and gender had fixed effects on the TF rates while diabetes status had a time-dependent effect. Applying the procedures proposed in the previous sections, we obtained $\hat{\gamma} = (0.0037, 0.1279)'$ with the estimated standard errors of 0.0221 and 0.0449 for age and gender, respectively. Fig. 2 presents the estimated effects of diabetes status, both the smooth estimate of the regression function $\beta(t)$ given by (7b) and the estimate of the cumulative regression function $B(t)$ along with its estimated 95% confidence bands. These results indicate that the male patients seem to have significant higher TF rate than the female patients, and in general the patients with diabetes seem to experience fewer TFs than those without diabetes at the start of CAPD treatment. More specifically, for the patients with diabetes at the start of CAPD treatment, the TF rate can be relatively higher initially and then start to go down after about three and half years. On the other hand, the TF rate did not seem to be related to the age of the patients at the start of CAPD treatment.

To further test the time-dependence and the significance of the effect of the diabetes status on the TF rate under the current model, we again carried out the test procedures based on $\mathcal{F}_1$ and $\mathcal{F}_2$ as well as $\mathcal{F}_3$. All three p-values were less than 0.0001. This confirmed that the diabetes status indeed had significant and time-varying effect on the rate of TF.

8. Concluding remarks

This paper has discussed regression analysis of multivariate recurrent event data when some covariates may have time-varying effects on recurrences of the events of interest with the focus on estimation of covariate effects. As commented before, recurrent event data occur in many fields, and a number of statistical approaches have been developed for situations where covariate effects are time-independent. In contrast, there exists little research on the case considered here, and there exist many situations in which some or all covariate effects could vary with time rather than being constants.

In the preceding sections, we have focused on estimation of the cumulative regression functions $B_0(t)$. As mentioned above, sometimes one may be interested in the estimation of $\beta_0(t)$ as well as $\lambda_{0k}(t)$, and one approach is to apply kernel estimates. For these, it would be helpful to develop some data-based methods for the selection of the bandwidth, which is beyond the scope of this paper. In both the simulation study and the example reported in the previous sections, the bandwidth was selected manually to give reasonably smooth estimates.

As discussed before, one advantage of the proposed methodology is that it leaves the association among different types of recurrent event arbitrary since it is usually difficult to provide correct specification of the association. For most situations, this or model (1) is sufficient as long as the main goal is the estimation of covariate effects. If the study goal is an estimation of the association or prediction, on the other hand, one may have to specify the association structure or the complete structure of the underlying model.

Acknowledgments

The authors would like to thank the Editor, an Associate Editor and two reviewers for their constructive and insightful comments and suggestions that greatly improved the paper. The first author's research was partly supported by the National Natural Science Foundation of China Grants (Nos. 10731010 and 10721101) and the National Basic Research Program of China (973 Program) (No. 2007CB814902). The authors also thank the Canadian Institute for Health Information for providing the data used in this paper. However, the conclusions and statements given in the paper are purely those of the authors and not the Canadian Institute for Health Information.

Appendix A. Regularity conditions

In this appendix we will describe the regularity conditions needed for the asymptotic properties given above. For this, suppose that, for $k = 1, \ldots, K$, the limits of $S_k^{(0)}(t;\beta,\gamma)$, $S_k^{(x)}(t;\beta,\gamma)$, $S_k^{(z)}(t;\beta,\gamma)$, $E_k^{(x)}(t;\beta,\gamma)$, $E_k^{(z)}(t;\beta,\gamma)$, $E_{xx}^{(k)}(t;\beta,\gamma)$, $E_{zx}^{(k)}(t;\beta,\gamma)$ and $E_{zz}^{(k)}(t;\beta,\gamma)$ exist and let $s_k^{(0)}(t;\beta,\gamma)$, $s_k^{(x)}(t;\beta,\gamma)$, $s_k^{(z)}(t;\beta,\gamma)$, $e_k^{(x)}(t;\beta,\gamma)$, $e_k^{(z)}(t;\beta,\gamma)$, $e_{xx}^{(k)}(t;\beta,\gamma)$, $e_{zx}^{(k)}(t;\beta,\gamma)$ and $e_{zz}^{(k)}(t;\beta,\gamma)$ denote the corresponding limits. Furthermore, let $s_k^{(0)}(t)$, $s_k^{(x)}(t)$, $s_k^{(z)}(t)$, $e_k^{(x)}(t)$, $e_k^{(z)}(t)$, $e_{xx}^{(k)}(t)$, $e_{zx}^{(k)}(t)$ and $e_{zz}^{(k)}(t)$ denote these limits at $\beta(t) = \beta_0(t)$ and $\gamma = \gamma_0$, $\bar{z}_k(t)$ be the $n \times p$ matrix with rows $e_k^{(x)}(t)^T$ and $\bar{z}_k(t)$ the $n \times q$ matrix with rows $e_k^{(z)}(t)^T$. Define

$$d(\tau) = \sum_{k=1}^{K} \int_0^{\tau} \left[e_{zz}^{(k)}(t) - a_z(t) a_x(t)^{-1} e_{zx}^{(k)}(t)^T \right] \mathrm{d}\mu_{0k}(t),$$

$$a_z(t) = \sum_{k=1}^{K} e_{zx}^{(k)}(t) \lambda_{0k}(t),$$

and

$$a_x(t) = \sum_{k=1}^{K} e_{xx}^{(k)}(t) \lambda_{0k}(t).$$

The required regularity conditions are as following.

(C1) $\beta_0(t)$ and $\mu_{0k}(t)$ are three times continuously differentiable for $t \in [0, \tau]$.

(C2) $X_i(t)$ and $Z_i(t)$ are of bounded variation on $[0, \tau]$.

(C3) $d(\tau)$ and $a_x(t)$ are nonsingular for $t \in [0, \tau]$.

(C4) G is a symmetric and continuous kernel function with a compact support satisfying $\int G(u)\mathrm{d}u = 1$; $h = O(n^{-\alpha})$, where $1/8 < \alpha < 1/4$.

(C5) $s_k^{(j)}(t;\beta,\gamma)$ is uniformly continuous with respect to $(t,\beta,\gamma) \in [0,\tau] \times \mathcal{B} \times \Theta$, where $\mathcal{B}$ is a compact set of R^p that includes a neighborhood of $\beta_0(t)$ for $t \in [0, \tau]$, Θ is a compact set of R^q including γ_0, and j denotes 0, x or z.

Appendix B. Supplementary proof

Supplementary proof associated with this article can be found, in the online version [21], at doi:10.1016/j.jmva.2009.08.00.

References

[1] P.K. Andersen, O. Borgan, R.D. Gill, N. Keiding, Statistical Models Based on Counting Processes, Springer-Verlag, New York, 1993.
[2] R.J. Cook, J.F. Lawless, Interim monitoring of longitudinal comparative studies with recurrent event responses, Biometrics 52 (1996) 1311–1323.
[3] R.J. Cook, J.F. Lawless, The Analysis of Recurrent Event Data, Springer-Verlag, New York, 2007.
[4] R.J. Cook, J.F. Lawless, J.C. Nadeau, Robust tests for treatment comparisons based on recurrent event responses, Biometrics 52 (1996) 557–571.
[5] P.K. Andersen, R.D. Gill, Cox's regression model for counting processes: A large sample study, Ann. Statist. 10 (1982) 1100–1120.
[6] J.F. Lawless, C. Nadeau, Some simple robust methods for the analysis of recurrent events, Technometrics 37 (1995) 158–168.
[7] D. Lin, L. Wei, I. Yang, Z. Ying, Semiparametric regression for the mean and rate function of recurrent events, J. Roy. Statist. Soc. Ser. B 69 (2000) 711–730.
[8] D. Lin, L. Wei, Z. Ying, Semiparametric transformation models for point processes, J. Amer. Statist. Assoc. 96 (2001) 620–628.
[9] R.L. Prentice, B.J. Williams, A.V. Peterson, On the regression analysis of multivariate failure time data, Biometrika 68 (1981) 373–379.
[10] H. Abu-Libdeh, B.W. Turnbull, L.C. Clark, Analysis of multi-type recurrent events in longitudinal studies: application to a skin cancer prevention trial, Biometrics 46 (1990) 1017–1034.
[11] J. Cai, D.E. Schaubel, Marginal means/rates models for multiple type recurrent event data, Lifetime Data Anal. 10 (2004) 121–138.
[12] L. Clegg, J. Cai, P.K. Sen, A marginal mixed baseline hazards model for multivariate failure time data, Biometrics 55 (1999) 805–812.
[13] D.E. Schaubel, J. Cai, Rate/mean regression for multiple-sequence recurrent event data with missing event category, Scand. J. Statist. 33 (2006) 191–207.
[14] C.F. Spiekerman, D. Lin, Marginal regression models for multivariate failure time data, J. Amer. Statist. Assoc. 93 (1998) 1164–1175.
[15] Z. Cai, Y. Sun, Local linear estimation for time-dependent coefficients in Cox's regression models, Scand. J. Statist. 30 (2003) 93–111.
[16] D. Lin, Z. Ying, Semiparametric and nonparametric regression analysis of longitudinal data (discussion), J. Amer. Statist. Assoc. 96 (2001) 103–126.
[17] T. Martinussen, T.H. Scheike, A semiparametric additive regression model for longitudinal data, Biometrika 86 (1999) 691–702.
[18] T. Martinussen, T.H. Scheike, I.M. Skovgaard, Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models, Scand. J. Statist. 29 (2002) 57–74.
[19] T.H. Scheike, T. Martinussen, On estimation and tests of time-varying effects in the proportional hazards model, Scand. J. Statist. 31 (2004) 51–62.
[20] K. Liang, S.L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986) 13–22.
[21] L. Sun, L. Zhu, J. Sun, Regression analysis of multivariate recurrent event data with time-varying covariate effects, Technical report, Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences. It is available from http://english.amt.ac.cn/sunliuquan.